
Parcours interprétatifs (inter)textuels : vers une assistance informatique

Computer-aided interpretation of texts and intertexts

Ludovic Tanguy et Théodore Thlivit



Édition électronique

URL : <http://journals.openedition.org/praxematique/1999>

DOI : 10.4000/praxematique.1999

ISSN : 2111-5044

Éditeur

Presses universitaires de la Méditerranée

Édition imprimée

Date de publication : 10 janvier 1999

Pagination : 185-215

ISSN : 0765-4944

Référence électronique

Ludovic Tanguy et Théodore Thlivit, « Parcours interprétatifs (inter)textuels : vers une assistance informatique », *Cahiers de praxématique* [En ligne], 33 | 1999, document 6, mis en ligne le 01 janvier 2014, consulté le 08 septembre 2020. URL : <http://journals.openedition.org/praxematique/1999> ; DOI : <https://doi.org/10.4000/praxematique.1999>

Tous droits réservés

Ludovic TANGUY

(ISSCO – Université de Genève/ERSS – Université de Toulouse Le Mirail).

Théodore THLIVITIS

(IASC – Ecole Nationale Supérieure des Télécommunications de Bretagne).

Ludovic.Tanguy@univ-tlse2.fr - Theodore.Thlivitis@enst-bretagne.fr

Parcours interprétatifs (inter) textuels : vers une assistance informatique

1. Introduction

Cet article reprend et synthétise un ensemble de travaux de formalisation des principes sémantiques liés à la textualité (Tanguy 97) et à l'intertextualité (Thlivitis 98). Sans insister sur les aspects formels de ces travaux, nous présentons surtout les possibilités et les conditions de leur utilisation dans une approche de compréhension de textes assistée par ordinateur.

Nous tâcherons dans un premier temps de définir les particularités de l'assistance informatique que nous proposons comme alternative à une automatisation, notamment par la prise en compte de la subjectivité d'une lecture. Nous aborderons ensuite l'outillage descriptif pour la représentation informatique, inspiré de la sémantique interprétative (Rastier 87), et son extension à certaines formes génériques de l'intertextualité. Nous envisagerons dès lors l'intertexte à la fois comme une forme de justification des attributions de sens à un texte, mais aussi comme le lieu privilégié de l'acte interprétatif lui-même.

1.1. Textes et informatique

Nous préciserons tout d'abord les conditions générales de notre approche informatique, en justifiant par rapport à celle-ci nos choix

quant aux types de représentation sémantique, et le cadre général de la place de l'automatisation dans cette approche.

1.1.1. Formalisme de représentation du sens

Notre paradigme est celui de la sémantique différentielle. Outre les considérations théoriques sous-jacentes sur la nature du langage, cette approche est surtout plus souple dans sa manipulation qu'une conception référentielle, où la partialité des représentations sémantiques est plus discutable, et la dépendance d'une description du monde de référence, ou d'une ontologie est limitatrice. De plus, elle est par nature plus apte à se mettre en place dans un rapport proprement intertextuel, en ce sens que les unités de signification sont ancrées sur des relations entre signifiés : mots, lexies ou textes.

Il est intéressant d'ailleurs de noter la popularité croissante de ce type d'approches dans le cadre du traitement automatique de la langue appliqué (notamment) aux corpus. La simplicité des descriptions permet leur application économique à des textes d'une grande variété, et la notion centrale de classe sémantique apporte un mode d'exploitation immédiat aux méthodes statistiques de classification des données lexicales en contexte, comme dans (Assadi 98, Pichon & Sébillot 99).

La micro-sémantique différentielle offre de plus une facilité d'implémentation informatique, et permet du point de vue de la formalisation d'aborder la description graduellement. Depuis les simples classes lexicales décrites sémantiquement, il est possible, comme on le verra, de complexifier l'organisation paradigmatique, et bien entendu, de capter des considérations sémantiques à l'échelle du texte et de l'intertexte.

1.1.2. Homme, machine et anthropocentrisme

Les notions que nous mettrons en place, ou du moins pour lesquelles nous préparons le terrain, sont de celles que l'automatisation n'a fait qu'effleurer. Le postulat est ici l'inséparabilité du sens et de l'individu interprétant : le système informatique sera donc conçu entièrement dans un esprit d'assistance et d'accompagnement d'une vision subjective.

La question ici est alors celle de la *justification* des attributions de sens :

1. Une première approche est celle de la liberté totale laissée à l'individu, sans contrainte d'aucune sorte. Si elle répond à une certaine vision du sens d'un texte, reconnaissant le droit du lecteur à n'y projeter que des connaissances personnelles, elle n'en reste pas moins stérile, voire inutilisable.

2. Une deuxième possibilité est celle d'une forme de cohérence interne d'un acte interprétatif. Cette cohérence peut prendre la forme d'un ensemble minimal de contraintes formelles, et surtout, ici, d'un appel à l'intertexte comme degré nécessaire de la justification. En plus des principes herméneutiques et philologiques de la théorie de l'intertexte, il peut devenir, dans son usage, similaire à une source externe de référence pour une attribution de caractéristiques linguistiques à un élément en contexte. À ce propos nous pouvons citer Paul Ricœur : « [...] à la faveur de cette oblitération du rapport au monde, chaque texte est libre d'entrer en rapport avec tous les textes qui viennent prendre la place de la réalité circonstancielle montrée par la parole vivante. Ce rapport de texte à texte, dans l'effacement du monde sur quoi on parle, engendre le quasi-monde des textes ou *littérature* » (Ricœur 86, p. 141). Par cette prise en compte, l'acte interprétatif auquel nous souhaitons apporter un mode de description se définit comme l'inscription d'un texte dans un contexte culturel, dont une manifestation opérationnelle est un ensemble de textes.

On discerne donc ici le rôle réservé à la machine. En lieu et place d'un automatisme calculatoire d'attribution de marqueurs sémantiques, il s'agit avant tout d'un travail d'*administration*, de vérification limitée et bien sûr de *représentation* du sens exprimé par l'utilisateur. Un dernier avantage de l'assistance informatique est celui de la *suggestion* et de la *motivation*. Dans le cadre de l'exploration assistée, ou de la didactique, nous verrons comment certaines contraintes formelles, comme celles présentées dans (Tanguy 97) peuvent servir de base à un dialogue entre l'homme et son outil dans la construction d'une représentation sémantique plus large et plus complète. De plus, une *visualisation* adéquate d'une représentation du sens exprimé par le lecteur peut faciliter la *réutilisation* des résultats interprétatifs par différents lecteurs, dans différentes interprétations et à travers différents textes.

1.2. La notion d'interprétation

1.2.1. Cadre général et définitions

Nous utiliserons ici une définition de *l'interprétation* comme un acte d'assignation de sens orientée par un objectif, organisée le long d'un parcours, et située par rapport à un individu et un ensemble de contextes.

De même nous définissons la *lecture* comme le produit d'une interprétation, un produit qui représente les relations sémantiques et les assignations de sens de l'acte d'interprétation, mais aussi les contextes textuels et intertextuels d'où sont issues ces assignations.

Concrètement, cette assignation de sens se traduit par l'association d'un ensemble de marqueurs à des unités linguistiques. Le premier niveau d'attribution est celui d'un marqueur sémantique, ou *sème*, comme constituant d'un signifié correspondant à une unité lexicale simple ou complexe : le *sémème*. Cette terminologie reprend les définitions de (Rastier 87).

Il est important, à ce stade, de marquer la distinction de cette méthode avec son utilisation classique en traitement automatique de la langue. La notion d'étiquetage sémantique, qui constitue dans le traitement de corpus électronique, un correspondant à l'étiquetage morpho-syntaxique, s'opère généralement à partir d'un ensemble fini de marqueurs, répondant à une organisation avant même d'être utilisés pour décrire un texte particulier. Nous verrons dans notre approche la porte s'ouvrir à un ensemble non restreint de marqueurs sémantiques. D'autre part, l'étiquetage sémantique se fait généralement à partir d'une description ontologique préalable, organisation a priori stable des signifiés ou des concepts repérés dans un corpus. De telles méthodes projectives sont ensuite utilisables pour des opérations de désambiguïsation ou de caractérisation sémantique de sous-parties de corpus.

Dans tous ces cas, le caractère fini du corpus et des unités d'étiquetage est requis.

Nous utiliserons au contraire non seulement un ensemble ouvert de descriptions, mais surtout un mode d'organisation souple de ces unités descriptives. Si ces descriptions ne visent pas « le sens » du texte analysé, mais une certaine vision partielle, elles n'en constituent pas moins un type de ressources réutilisables.

1.2.2. Intertexte, justification interprétative et contraintes

Une première conséquence de cette ouverture est la *limitation* et la *mise en contexte* des descriptions ainsi mises en place. L'immersion d'un texte dans un univers interprétatif ouvert dans plusieurs directions ne peut espérer atteindre une exhaustivité descriptive. Il est important, dès lors, de préciser les dimensions de cette limitation.

1. *Les unités décrites* : l'intégralité des unités lexicales, et même des textes du corpus ne sera pas décrite à l'aide de marqueurs sémantiques. Toute description se concentre a priori sur un ensemble de phénomènes textuels restreints. Cette limitation ne concerne cependant pas la nature ou la taille des unités décrites, et nous verrons que le formalisme de représentation est au contraire d'une souplesse suffisante pour aborder de façon uniforme les différents paliers de la textualité et de l'intertextualité.

2. *L'étendue et la couverture des descriptions* : Au contraire d'une approche référentielle, qui chercherait à l'aide de représentations sémantiques à calquer une réalité ontologique externe aux unités textuelles, le type de description est ici ouvert. Nos inspirations théoriques en sémantique différentielle tendent d'ailleurs à se concentrer sur un palier linguistique présumé autonome dans un premier temps, donnant aux marqueurs sémantiques utilisés le statut d'indices de l'organisation du système de la langue. Dès lors, il serait vain d'attendre de l'ensemble des sèmes associés à une unité lexicale une couverture équivalente à celle de la lexicologie.

3. *La subjectivité et la justification* : comme nous y faisons allusion précédemment, les approches de description sémantique que nous envisageons ici ne comptent pas échapper à la subjectivité de l'acte interprétatif. Ce que nous proposons pour l'accompagner est une certaine forme de rationalité des déclarations sémantiques. L'attribution d'un sème à un signifié ne peut être entièrement justifiée en l'absence d'un ensemble exhaustif (et donc utopique) de ressources et de références. Dans l'absolu, les sèmes sont les organisateurs minimaux et suffisants du système d'une langue. Une réduction de l'étendue du système leur enlève leur justification. Mais, tout comme la linguistique descriptive fait appel à des textes et des exemples qualifiés d'attestés, une forme de justification peut apparaître par l'introduction d'autres textes, au sein

desquels l'attribution sémique peut être explicitée directement, notamment par le biais de cooccurrence ou de structures syntaxiques explicites. Dans cette optique, certains textes au statut particulier, comme les ouvrages de références ou les encyclopédies, jouent naturellement ce rôle, mais ne doivent pas en garder l'exclusivité.

Dès lors, on peut entrevoir le double statut de l'intertexte dans cette optique. Comme objet d'étude à part entière, car il constitue le lieu minimal de certaines études comparatives de la textualité : il est alors plus proche du corpus dans sa nature. Il consiste en un ensemble de données, que l'interprète confronte à des hypothèses, et doit alors être abordé par un ensemble de concepts propre à décrire les différentes relations entre textes, en ce qu'elles diffèrent des structures syntagmatiques intratextuelles. Mais il prend également la forme d'un objet construit, d'un matériau expliciteur rassemblé dans un objectif qui, lui, peut se limiter à un seul texte.

Dans les deux cas, les mécanismes d'attribution de sens, et en particulier des phénomènes comme celui de *l'afférence*, peuvent prendre une ampleur et une précision formelle dans un cadre intertextuel. Nous verrons dans la suite quelques outils descriptifs simples pour décrire l'apport de notions et d'associations d'un texte à l'autre, ou d'un ensemble de textes vers lui-même.

Toutefois, la prise en compte de cet univers supérieur au texte ne doit pas éclipser le jeu complexe des relations sémantiques locales, ne pouvant être séparées de la notion de textualité. Les mécanismes descriptifs que nous proposons ici sont donc envisagés dans une dynamique qui fait évoluer la granularité du texte à son ensemble, et ce dans les deux directions. Chaque passage d'un niveau à l'autre aura toujours pour objectif l'enrichissement de la structure sémantique des deux. La continuité, on le verra, sera principalement assurée par des notions comme celle de *classe sémantique* et d'*isotopie*, enrichies par les différentes relations entre les textes au sein desquels elles évoluent.

2. Eléments formels de la sémantique interprétative

Nous résumons ici les principaux concepts descriptifs de la Sémantique Interprétative de F. Rastier, notamment tels qu'ils sont décrits

dans (Rastier 87), en en résumant les approches formelles proposées dans (Tanguy 97).

2.1. L'isotopie

La notion d'isotopie est le lieu qui permet de relier des notions sémantiques de l'échelle du texte à celles des unités qui le composent. Définie comme récurrence d'un sème le long de l'axe syntagmatique, elle est donc, suivant cette direction du global vers le local, une explicitation de notions telles que le thème ou le champ sémantique, depuis la généralité de l'ensemble vers la particularité des parties.

Le premier passage consiste en cette projection vers les unités locales (dont la forme la plus générale est celle de la lexie, définie chez Pottier), elle correspond à l'explication de « Ce texte parle de X » par « Ce texte contient des expressions qui traduisent la notion de X ». La liste des unités locales ainsi repérées dans un texte permet donc une première analyse, et une forme d'argumentation de cette notion générale.

Une isotopie est donc la relation qui lie un sème à un ensemble de sémèmes présents dans un texte. Elle forme dès lors une première notion de classe sémantique, puisqu'elle définit une relation d'équivalence entre ces sémèmes. Mais elle est également construite sur d'autres formes d'organisation des signifiés.

2.2. Généricité et spécificité

F. Rastier distingue deux principaux types de relation en langue entre un sémème et un sème, donc deux types de sèmes : *générique* et *spécifique*. Le premier traduit une notion de classe sémantique constituée de sémèmes proches, relevant d'une même organisation paradigmatique. La nature de ces classes, leur nombre et leurs statuts relatifs peut être discuté, mais nous ne retenons ici que leur principe général. Une classe sémantique est un ensemble de sémèmes d'un même texte décrits par un seul et même sème générique. Nous assimilerons dans la suite la notion de classe sémantique à celle de *taxème*, ou classe sémantique minimale qui joue, dans l'appareil descriptif de la Sémantique Interprétative, le rôle d'une unité organisatrice nécessaire, par rapport à d'autres types de classes plus générales (comme les domaines et les dimensions).

Cette relation de genericité possède bien sûr une relation duale de spécificité, ou de différenciation entre deux signifiés appartenant à une même classe, mettant ainsi en place la notion de sème spécifique. Nous pouvons formaliser cette notion par celle d'une opposition binaire entre deux sèmes d'un même taxème. C'est alors cette opposition qui est décrite par un sème, plus que le sémème, de même qu'un sème générique est attribué à une classe, et non à ses éléments. Par exemple, si nous voulons représenter ainsi une description sémantique naïve de certains noms d'animaux présents dans *Les Chants de Maldoror* (Lautréamont), nous pouvons le faire à l'aide d'une classe sémantique, ou taxème, à laquelle nous attribuons le sème /Animal/, et en mettant en place un réseau d'oppositions binaires entre les éléments de cette classe, ces oppositions étant également qualifiées par des sèmes, comme présenté dans la Figure 1 :

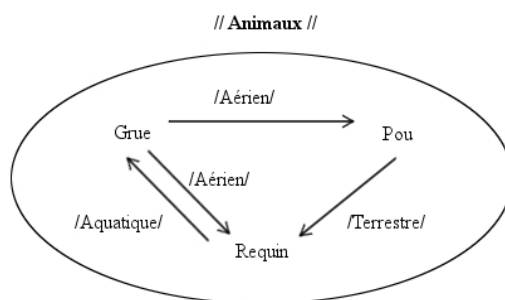


Figure 1 : Taxème et sèmes spécifiques

Dans cette figure, chaque sémème de la classe se voit donc qualifié par un ou plusieurs sèmes spécifiques, par opposition à d'autres sémèmes de la même classe.

Cette première ébauche de formalisation permet donc, en plus de traduire l'association directe de sèmes à des sémèmes, de proposer cette même association au travers d'un ensemble de relations entre sémèmes, correspondant à un niveau d'organisation sémantique que les sèmes viennent simplement qualifier. Cette structuration des attributions sémiqes permet déjà une forme d'assistance à l'exploration d'une

certaine forme de la textualité : comme il a été proposé dans (Tanguy 97), elle permet de mettre en place, avec l'aide d'un outil informatique, PASTEL (Tanguy & Thlivitis 96), un jeu interprétatif dans lequel l'utilisateur déclare initialement un certain nombre d'isotopies sur un texte, et poursuit l'explicitation de son interprétation en construisant une structure dans laquelle les unités lexicales entretiennent des relations des deux types que nous venons de décrire.

Le rôle de la machine, dans ce cas, est d'imposer une forme d'organisation des relations sémiques afin que la structure construite autour de la vision de l'utilisateur réponde à un certain nombre de contraintes formelles. Un exemple de ces contraintes est la nécessité, pour chaque taxème, considéré sous l'angle des relations de spécification qu'entretiennent les sèmes qu'il contient, de posséder une structure de graphe suffisamment complète pour que chaque élément de la classe soit complètement spécifié, et possède ainsi une identité sémantique. La mise en place de cet outil constitue donc un dialogue entre l'interprète et le programme, par un jeu de questions posées par la machine à l'utilisateur, dont les réponses sont autant de sèmes qui viennent enrichir la structure sémantique. Ce dialogue peut amener l'interprète à envisager de nouvelles relations sémantiques qu'il aurait initialement négligées, et à découvrir par exemple de nouvelles thématiques transversales entre les principaux thèmes repérés initialement.

Cette explicitation d'un ensemble initial d'intuitions de la part de l'interprète peut également être vue comme une forme de justification de celles-ci : la construction d'une structure sémantique de ce type met en quelque sorte à l'épreuve les isotopies principales du texte, et vient les renforcer et les enrichir.

2.3. Inhérence et afférence

Cependant, ce type d'explicitation des attributions de sèmes, si elle permet d'asseoir la présence d'une thématique dans un texte en multipliant les relations et les unités qui la manifestent, ne permet pas pour autant de justifier ces attributions qui, rappelons-le, n'engagent que le seul utilisateur qui les a déclarées (le programme cité précédemment ne faisant appel à aucune base de données générique de sèmes). Si certaines notions font appel à un niveau de systématisme du langage que

l'on ne peut remettre en cause (comme les domaines naturels des animaux dans l'exemple précédent), certaines notions sont, elles, moins acceptables, et leur justification doit généralement sortir du texte, sont pour autant revenir à une définition du type encyclopédique. C'est le cas, par exemple, des connotations négatives ou agressives du bestiaire de Lautréamont, faisant ici appel à une norme sociale ou culturelle.

C'est cette distinction qu'explique Rastier en définissant une nouvelle distinction dans les types de sèmes : *inhérents* et *afférents*. Les premiers relèvent, comme nous venons de le rappeler, d'un niveau de langue général, alors que les sèmes afférents proviennent de normes locales (sociales, voire individuelles), ou encore du contexte. Dans ce cas, et pour continuer à jouer le jeu de la construction d'une structure formelle explicatrice d'une interprétation subjective, cette notion d'afférence doit, elle aussi, se voir attribuer un support duquel elle peut recevoir une forme de rationalité.

Par opposition à l'inhérence, l'afférence est donc censée prendre en compte la notion de contexte (*e.g.* des éléments provenant du texte, d'un autre texte, ou d'un genre) et c'est exactement cette mise en contexte opérée par un individu ayant des compétences interprétatives propres qui la rend aussi importante dans toute interprétation et en même temps problématique, contestable et, de par sa subjectivité, informatiquement inutilisable. Son importance apparaît aussi par la position des afférences au cœur des choix des parcours interprétatifs du lecteur. Et, de leur statut contestable témoignent les champs de bataille entre différentes visions interprétatives¹.

Pour donner un exemple (limité mais concret) des difficultés que pose la notion d'afférence, nous examinons une lecture du sonnet « Salut » de Mallarmé effectuée dans (Rastier 89). À un moment de son analyse, Rastier met en avant la récurrence du trait sémantique (*sème*) /navigation/ sur un ensemble de sémèmes du sonnet (*i.e.* 'Salut', 'écume', 'sirènes', 'se noie', etc.). Cette récurrence se décline selon trois types d'attribution :

1. Le sème /navigation/ est attribué en tant que sème *inhérent* du sémème en question. Par exemple 'naviguons' ou 'poupe' reçoivent le sème /navigation/ comme définitoire, hors contexte. L'inhérence relève,

¹ Cf. la polémique relative au sonnet de Mallarmé dans (Rastier 89, p. 226).

dans la définition qu'en fait Rastier, d'une stabilité correspondant aux paliers supérieurs du système linguistique, ce qui est bien le cas ici, comme nous le confirment les définitions encyclopédiques de ces termes.

2. Le sème /navigation/ est attribué en tant que sème *afférent* à 'sirènes' ou à 'écume'. Autre exemple, l'attribution de /navigation/ à 'Salut', avec le commentaire « port de salut, ancre de salut » (p. 226). Cette première notion d'afférence correspond donc à un niveau de systématique moindre que pour les attributions précédentes, en ce sens que le sème ici attribué n'est pas définitoire, mais tout de même justifié dans un contexte large, celui d'une norme sociolectale.

3. Le sème /navigation/ est enfin attribué en tant que sème *afférent en contexte* à 'solitude' avec le commentaire « solitude sur la mer (lieu désert) » ou à 'souci' avec le commentaire « préoccupation des marins, objectif de la navigation » et avec un renvoi vers la page 72 des *Œuvres complètes* de Mallarmé.

Tout comme la notion d'isotopie permet un début d'explicitation de la notion de thématique et des structures sémantiques du texte en construisant à partir de sa constatation un mode d'organisation paradigmatique des signifiés, les attributions afférentes de sèmes doivent elles aussi pouvoir bénéficier d'un mode d'explicitation.

L'idée de départ est la suivante : une *attribution sémique* (et notamment une *afférence* en contexte) est contrainte d'être fondée sur une *interprétation* précédemment établie dans un *contexte proche*.

Ceci veut dire que l'attribution du sème /navigation/ à 'solitude' dans le contexte du sonnet de Mallarmé demande une interprétation préalable qui puisse *attester* une telle attribution dans un *contexte proche*. Nous appelons ces interprétations préalables *interprétants* de la nouvelle interprétation puisqu'elles fournissent les attributions de sèmes nécessaires à sa constitution.

Concrètement, si le lecteur veut établir l'isotopie /navigation/ dans le sonnet de Mallarmé (cf. la Figure 2) il lui sera demandé de justifier les attributions du sème à chacun des sémèmes de l'isotopie en choisissant par exemple un ensemble d'interprétations qui explicitent ces attributions. Par exemple, dans la figure, l'*Interprétation*₁ établie dans le *Texte*₁ contient une attribution du sème /navigation/ sur le sémème

‘sirènes’. L’*Interprétation*₁, sélectionnée d’abord par le lecteur, est ensuite utilisée par le système pour créer une afférence du sème /navigation/ sur ‘sirènes’ dans la nouvelle interprétation dans *Salut*.

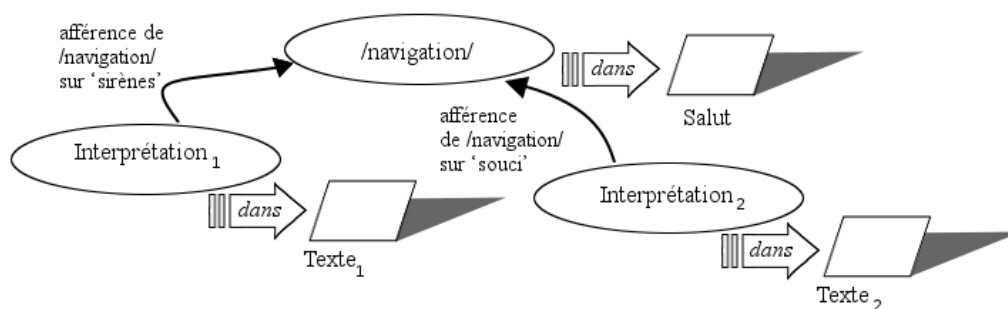


Figure 2 : L’afférence intertextuelle dans le parcours interprétatif

Le lieu de passage d’une attribution à l’autre nécessite bien entendu une lexie commune aux deux textes (‘solitude’ ou ‘sirènes’), et ce sont les sèmes attribués au sémème du texte source qui sont transposés vers celui du texte cible. De ce fait, une telle extension des attributions de sème revient à étendre la définition de l’isotopie. Là où elle était restreinte à la récurrence d’un sème le long de l’axe syntagmatique d’un texte, elle devient lien sémantique entre des unités présentes dans plusieurs textes. Les détails de ces relations et contraintes sont décrits dans la section 3.

L’afférence est alors vue en tant que *transformation* d’une *interprétation* préalable en *interprétant* pour la constitution d’une nouvelle classe sémantique. Autrement dit, selon ses intentions interprétatives, le lecteur peut réutiliser un résultat interprétatif préalable en tant qu’*interprétant* pour une nouvelle interprétation. L’afférence se situe ainsi au cœur d’une *réutilisation du sens*, certes contrainte, mais pouvant aller au-delà d’un texte, d’un lecteur et d’un objectif interprétatif.

Ce dépassement de la limite d'un texte fait que l'afférence est dans le cas général *intertextuelle*. Par exemple, la constitution de la classe sémantique /navigation/ de notre exemple demande la prise en compte d'au moins trois extraits des *Œuvres complètes* de Mallarmé. Cette prise en compte d'autres textes n'est pas marginale. Dans l'analyse de ce même sonnet, il faudra d'abord interpréter une note bibliographique rédigée par Mallarmé (notamment, « ce sonnet, en levant le verre, récemment, à un Banquet de *La Plume*, avec l'honneur d'y présider »²) pour pouvoir afférer, par exemple, les sèmes /mousse de champagne/ à 'écume' et /verre à champagne/ à 'coupe'.

C'est ainsi, en tant qu'élément fondamental pour la représentation d'un *parcours interprétatif*, que nous nous servons dans un premier temps de la notion de *l'intertexte*.

3. Anagnose : une lecture intertextuelle

Pour expliciter les relations établies au sein d'un intertexte et ainsi « situer » les étapes du parcours interprétatif du lecteur, nous introduisons une notion opérationnelle, appelée *anagnose* (Thlivitis 98).

3.1. Définition

Une *anagnose* correspond simplement à un espace de travail comprenant un ensemble de textes, qui délimite les frontières textuelles d'une lecture. D'un point de vue sémantique, une *anagnose* comprend la production interprétative d'un lecteur selon un objectif de lecture.

Globalement, une anagnose constitue une interprétation couplée à un intertexte.

Quelques précisions sont ici nécessaires :

1. Une anagnose n'est pas un corpus ni un genre de textes. Il est donc nécessaire de la distinguer de la notion de corpus de référence en TALN classique, où l'on prend un ensemble de textes, soit comme un échantillon représentatif d'un ensemble de productions textuelles à des fins de généralisation d'un ensemble de concepts ou d'outils, soit comme un ensemble cohérent couvrant un genre, généralement délimité par une certaine forme de pragmatique ou d'univers référentiel. Une

² (Rastier 89, p. 227).

anagnose peut contenir des textes appartenant à un ou plusieurs corpus. Mais surtout il s'agit d'un intertexte a priori *ouvert* ; les textes sont insérés par le lecteur au fur et à mesure que leur prise en compte devient nécessaire relativement aux interprétations à établir et aux objectifs régissant la construction de l'anagnose.

2. Une anagnose n'est pas prescrite par l'auteur mais *établie par le lecteur* : il s'agit d'une *création d'interprétation*. Le lecteur reste libre de créer une anagnose qui ne respecte pas les prescriptions doxales, voire même la pratique sociale dont fait partie le texte. Il pourra ainsi, s'il y trouve un intérêt, interpréter par exemple un texte platonicien sous la lumière de textes de philosophes contemporains. L'anagnose, sous le point de vue d'un « ensemble de textes » reflète directement ses choix d'interprétation de telle manière qu'un autre lecteur, se contentant de survoler les textes utilisés dans l'anagnose ainsi que les liens correspondant aux relations sémantiques entre textes, peut déjà avoir une idée du cadre interprétatif de cette lecture.

3. Un texte n'appartient donc pas à un intertexte unique et il n'en suggère pas un non plus. Autrement dit, l'anagnose n'est pas une propriété intrinsèque du texte. Un texte peut appartenir à autant d'anagnoses que de visions interprétatives déclarées utiles par les lecteurs.

4. Un texte hors anagnose peut être vu comme une lexie hors texte : on peut lui trouver une signification mais son *sens* sera précisé au sein d'une anagnose, parmi d'autres textes qui l'influencent et qui sont influencés par lui. En d'autres termes l'interprétation d'un texte nécessite l'établissement d'un intertexte. Encore plus, l'interprétation d'un texte est *déterminée* par le choix de l'intertexte. Mais pour choisir l'intertexte il est nécessaire d'interpréter le texte, ne serait-ce que pour identifier l'axe thématique de l'organisation de l'intertexte. L'intervention du lecteur transforme cette boucle en spirale. Ce sont ses présomptions qui établissent l'intertexte initial (qui peut être minimal, le texte seul) et c'est son orientation interprétative qui guide d'une part les choix interprétatifs et de l'autre l'évolution de l'intertexte.

Plus formellement, une anagnose est un ensemble complexe de relations construites selon un ensemble de contraintes. Les détails formels étant explicités dans (Thlivitis 98) nous nous limitons ici à décrire simplement et intuitivement ces relations et contraintes.

L'ensemble des *relations* établies au sein de l'anagnose à un instant donné du parcours interprétatif du lecteur caractérise un état stable de l'anagnose, mais peut-être pas son état final. Les *contraintes* concernent précisément le passage d'un état à l'autre.

Finalement, et pour représenter la dynamique du parcours interprétatif, l'anagnose contient aussi les informations qui ont servi au passage d'un état à l'état suivant. Ces informations instancient en quelque sorte une ou plusieurs contraintes. C'est par exemple le cas des informations explicitées par les liens entre une nouvelle interprétation (se trouvant dans un nouvel état) et ses interprétants, i.e. des interprétations présentes dans l'état précédent. En d'autres termes, la construction d'une interprétation nécessite la présence des interprétants (c'est la contrainte) et, en représentant les afférences de sèmes entre les interprétations préalables et la nouvelle interprétation, le système enregistre l'information qui a servi au passage d'un état à l'état suivant.

Pour simplifier la présentation des relations et des contraintes, nous la décomposons en deux parties. Nous discutons d'abord des relations et contraintes représentant un état donné, et ensuite des relations et contraintes de passage d'un état à l'état suivant, notamment les justifications par afférences intertextuelles.

3.2. Représentation d'une interprétation

Une anagnose a une structure *sémantique* interne. Elle est le lieu où sont établies les relations sémantiques textuelles et intertextuelles. La représentation des relations sémantiques au sein d'un texte et notamment des isotopies suit de près les définitions de la Sémantique Interprétative telles qu'elles ont été présentées dans la section 2.

Pour le cas d'une isotopie, et plus généralement d'une relation sémantique, définie dans un texte, il s'agit d'un ensemble de relations entre un (ou plusieurs) sème(s) et un ensemble de sémèmes repérés dans le texte. A ces relations s'ajoutent des contraintes : par exemple pour le cas d'une classe sémantique celle d'un ensemble de relations supplémentaires pour les sèmes spécifiques entre les sémèmes pris deux par deux.

Reste la représentation des relations intertextuelles, e.g. /parodie/ ou /commentaire/ (cf. (Genette 82) pour une typologie assez élaborée des

relations intertextuelles et la contribution de Ioannis Kanellos dans ce même numéro). Ce qui nous intéresse ici n'est pas une liste exhaustive de ces relations mais un cadre formel pour leur représentation.

La représentation statique d'une telle relation intertextuelle (par opposition à la représentation de la dynamique de sa création, présentée plus bas) est effectuée à l'aide d'une relation entre un sème (e.g. /parodie/) et un ensemble de textes appartenant à l'anagnose. Ce type de relations concerne les textes considérés dans leur globalité. Il y a aussi souvent le besoin d'exprimer des relations entre des parties de textes, parlant d'un même thème, acteur, ou en général ayant diverses relations sémantiques. Pour ce faire, nous définissons aussi un deuxième type de relations intertextuelles entre des classes sémantiques définies à l'intérieur des textes. Par exemple une isotopie³ du sème /navigation/ concernant des parties de deux textes de Mallarmé sera une relation entre les isotopies déjà établies au sein de ces textes.

Nous remarquons tout de suite que ce type de relations intertextuelles ne concernent pas directement les sémèmes des textes. Car, nous rappelons que selon la définition de la notion d'isotopie dans la *Sémantique Interprétative* (Rastier 87) les éléments d'une isotopie correspondent à des éléments d'un seul texte. Même si la constitution d'une isotopie peut exiger un parcours interprétatif concernant différents textes, nous maintenons cette contrainte en interdisant toute relation sémantique directe entre des sémèmes n'entretenant pas de rapports syntagmatiques d'aucune sorte.

Cette contrainte de localité préserve la séparation entre l'auteur (*producteur* de la structure intra-textuelle) et le lecteur (*producteur* de la structure intertextuelle et du parcours interprétatif).

A ces relations sémantiques s'ajoutent bien sûr les relations entre les unités textuelles : le texte contient les lexies identifiées par les lecteurs et l'anagnose contient les textes qui y sont ajoutés lors des parcours interprétatifs.

³ Le terme isotopie est ici utilisé pour simplifier le discours. Il s'agit plutôt d'une isotopie *intertextuelle*.

3.3. Représentation du parcours interprétatif intertextuel

Nous passons maintenant aux relations et les contraintes qui gèrent et représentent le *parcours* interprétatif et qui concernent les liens entre une nouvelle interprétation et sa « justification », c'est-à-dire les interprétations qui fournissent les attributions de sèmes aux sémèmes de la nouvelle interprétation.

Une telle relation est présentée sur la Figure 3. Trois unités y participent : la nouvelle interprétation, un ensemble d'interprétations existantes et, pour chacune de ces interprétations, la partie sémantique qui est utile à la nouvelle interprétation. Par exemple, dans la figure, cette partie correspond à l'attribution du sème s_I à la lexie l_I (représentée par le rectangle de la figure). En effet, cette « partie utile » représente le « début » d'une afférence du sème s_I au sémème correspondant de la nouvelle interprétation.

Une afférence est donc une relation entre une attribution de sème réalisée et validée dans un texte-source, et une attribution du même sème à une autre occurrence du sémème dans le texte-cible.

L'explicitation de ces trois unités a donc une composante dynamique, et la mise en place de ces différentes relations est échelonnée dans le temps, le long du parcours interprétatif. La Figure 3 commente un tel parcours. L'isotopie du sème s_I concerne les lexies (sémèmes) l_1 , l_2 et l_3 d'un texte. Trois afférences de ce sème sont utilisées pour établir cette isotopie. Dans la figure, l'afférence sur l_I est issue d'une attribution du sème s_I dans une autre classe ou isotopie d'un autre texte de l'anagnose sur une autre occurrence de cette lexie (l_I et l_I sont équivalentes et ne diffèrent donc que par le lieu de leurs occurrences).

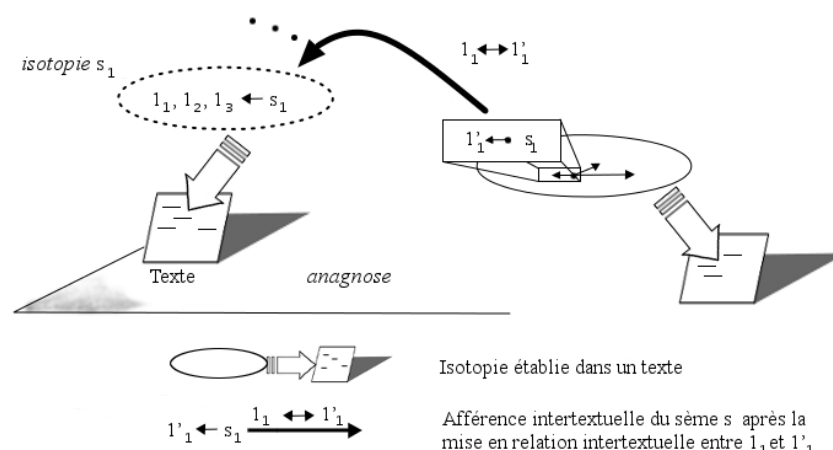


Figure 3 : Schéma abstrait de *parcours interprétatif* pour la création d'une classe sémantique isotopique.

Une instantiation possible de ce schéma pourrait être la suivante :

$s_1 = \text{'férocité'}$, $l_1 = l'_1 = \text{'pou'}$, $l_2 = \text{'requin'}$, ...

En d'autres termes, pour représenter les étapes d'un parcours interprétatif nous exigeons que le lecteur participe à un *jeu de construction* de classes sémantiques. Une classe sémantique représente une interprétation et la notion de *construction* permet de capter la partie dynamique du parcours interprétatif, notamment le passage d'une présomption d'interprétation à sa justification (dans une direction *de haut en bas*) ainsi que son inverse, le passage d'un ensemble d'interprétations à une synthèse (dans une direction *de bas en haut*).

Une question que l'on peut se poser naturellement concerne *l'étendue* de cette *justification* des interprétations. Autrement dit, qu'est-ce qui peut être considéré comme une *interprétation donnée*,

sans besoin de justification, pour que les processus de construction puissent ensuite être entamés ?

Deux possibilités sont prévues pour cette amorce :

1. l'interprétation est complètement explicitée au sein d'un texte. En effet, une *cooccurrence* de la lexie et de son sème permet une attribution simple du sème à un sémème. Tel est le cas par excellence des informations sémantiques tirées des dictionnaires, ou bien de certaines constructions syntaxiques⁴.

2. l'interprétation constitue une hypothèse de travail, centrale dans une lecture. Ces déclarations axiomatiques sont alors déclarées telles quelles dans l'anagnose en question, afin de bien expliciter la cohérence globale de la lecture. Plus tard, le lecteur peut décider de justifier ces hypothèses en élargissant son parcours interprétatif et en continuant son jeu de construction.

3.4. Commentaires

Dans le schéma de la Figure 3, la direction des flèches depuis la source sémique vers la classe cible suit une logique due au formalisme plus qu'à un processus interprétatif. Nous devons donc plutôt parler d'un parcours interprétatif *formel* par opposition au *parcours interprétatif réel* qui correspond à l'interaction de l'utilisateur avec le système.

La même représentation peut ainsi être instanciée selon deux directions opposées d'interaction avec l'utilisateur. Une première direction est celle qui va de la *source* vers le *résultat*. Par exemple, selon cette voie, dans le cas de la Figure 3, le lecteur identifie d'abord les interprétations de 'pou' qui font que le sème /férocité/ lui soit attribué explicitement, par exemple lorsque la notion de /férocité/ apparaît directement dans une lexie, « ... *féroce* comme un *pou* rouge » (P. Claudel). Ensuite, le lecteur utilise ces attributions comme afférences au sein des 'Chants' où il peut construire son isotopie de /férocité/.

⁴ Il est important de noter toutefois que nos propositions diffèrent de certaines approches en analyse des cooccurrences dans les corpus, puisque cette métamorphose d'une lexie cooccurrence en sème n'est ni systématique ni nécessairement directe. Une expression comme « laid comme un pou » permet d'affecter un sème plus général de /négativité/, et non seulement de /laideur/, une fois la relation entre /négativité/ et /laideur/ soit établie dans la lecture actuelle.

Cependant il existe la direction inverse, depuis la *présomption* de l'isotopie *s_I* (e.g. /férocité/) vers sa mise en exergue à l'aide des interprétations sources. Ceci voudrait dire, pour la Figure 3 que le lecteur établit d'abord l'isotopie /férocité/ dans les Chants en utilisant des attributions *ad hoc* du sème /férocité/. Ensuite, il cherche *pourquoi* et surtout *où* chacune de ces afférences peut être explicitée. C'est ainsi que le lecteur aura recours à d'autres textes qu'il lui faudra éventuellement interpréter pour établir ces interprétations sources qui fournissent l'attribution du sème /férocité/, par exemple en tant que sème spécifique ou générique au sein d'un taxème local.

Pour remplir cette tâche le lecteur peut demander l'aide du système, que ce soit par exploration d'un ensemble d'interprétations déjà effectuées, ou par recherche d'occurrences dans une base de textes. Le lecteur peut ainsi décrire quelques aspects de l'état souhaité et le système peut calculer et suggérer quelques possibilités. Par exemple le lecteur peut définir le sème et les éléments d'une nouvelle isotopie et demander au système les interprétants les plus proches dans l'anagnose.

Une dernière remarque concerne la fin de la constitution d'une anagnose et de ses relations. Il n'y a aucune contrainte formelle sur l'étendue valide d'une anagnose. C'est le lecteur qui décide du moment où l'intertexte est suffisamment étendu et l'anagnose suffisamment riche en relations sémantiques. Il pourrait effectivement continuer ses parcours interprétatifs en rajoutant de nouvelles relations et de nouveaux textes jusqu'à épuisement. Cependant, une anagnose, comme d'ailleurs tout texte, doit être considérée comme le produit d'un objectif interprétatif précis, limité et donc fini.

Ce jeu de construction de classes sémantiques obéit à des règles précises et l'on peut se demander si la liberté interprétative et l'originalité du lecteur se trouvent ainsi limitées. En effet ces règles sont issues de la théorie de la *Sémantique Interprétative* étendue à une *Sémantique Interprétative Intertextuelle* (Thlivitis 98) ; elles sont simples et intuitives, du moins dans le cadre d'une sémantique différentielle. Si les lecteurs doivent procéder d'une interprétation à l'autre de manière contrôlée, leur liberté est présente dans le choix de l'intertexte, dans le choix des parties textuelles sur lesquelles vont opérer les règles de

construction et dans le choix de la combinaison des constructions partielles vers les constructions finales.

En revanche, l'avantage de la présence de ces règles est qu'elles offrent une base commune qui facilite la consultation, la comparaison et la réutilisation des résultats interprétatifs.

À l'aide de ces outils minimaux, nous pouvons d'ores et déjà proposer une mise en place dynamique de certains principes interprétatifs.

4. Assistance Informatique

Nous reprenons ici, plus précisément, les grandes lignes des fonctions attribuées au programme informatique dans la construction d'un parcours interprétatif.

4.1. Rôle général de l'outil

La fonction principale est celle de la gestion d'une base de données mettant en jeu un schéma complexe construit autour de cinq principaux types de données : les *anagnoses*, les *textes*, les *lexies*, les *sèmes* et les *classes sémantiques*.

Sur ces types de données sont définies, comme nous avons vu au long de cette présentation, différents types de relations :

- Σ Relations de composition (textes – lexies, anagnoses – textes)
- Σ Relations ordonnées (entre textes et entre lexies d'un même texte)
- Σ Relations sémantiques (d'un sème à un sémème ou un texte). La notion d'isotopie est ici décomposée en un ensemble de telles relations sème-sémème, à condition que les sémèmes apparaissent dans un même texte.
- Σ Relations entre relations sémantiques (entre relations sémantiques de différents textes, ou entre une relation sémantique et les relations sémantiques qui ont servi à sa construction)

Cette base, dont la complexité apparaît déjà, doit être manipulée par un ensemble de fonctionnalités propres aux outils classiques de gestion : la saisie, le stockage, l'exploration et la représentation de ces données.

De plus, un ensemble de *contraintes formelles* sont déclarées au niveau de la constitution et de l'organisation de ces relations. Nous en avons esquissées certaines, dont l'activation a pour but d'inciter l'utilisateur à enrichir la base, et dont la gestion se doit elle aussi d'être informatisée.

4.2. Interface de manipulation et de visualisation générale

Le rôle de cette interface utilisateur est celui de la gestion d'une base de textes et des informations qui s'y rattachent. Elle propose donc à l'utilisateur des fonctionnalités d'organisation des différents textes, de sélection des sous-parties (lexies) de ceux-ci pour en constituer des éléments signifiants, de caractérisation de ces sous-parties, de création de nouvelles unités sémantiques s'intégrant dans le schéma général du parcours interprétatif, de suggestion de classes pouvant servir d'interprétant aux nouvelles classes en constitution, etc. Ces fonctionnalités sont accompagnées d'outils de recherche d'occurrences et cooccurrences, de visualisation multiple des données selon différents points de vue, etc.

Cependant, plus qu'un ensemble de fonctionnalités, l'outil informatique offre aux lecteurs un protocole d'interaction qui les accompagne dans leurs parcours interprétatifs car elle intègre déjà les contraintes présentées plus haut. A travers cette interaction le système invite le lecteur à participer à un « jeu » d'interprétation, dont les règles sont intégrées aux fonctionnalités du système.

A titre d'exemple nous présentons une partie de l'interface utilisateur du système dans la figure 4. Sur la partie gauche l'interface propose plusieurs vues sur la même représentation générale sous-jacente. On peut voir (en haut) deux vues générales sur les relations sémantiques, l'une sous forme de tableau et l'autre sous forme d'arbre de parcours interprétatifs, et (en bas) deux vues aussi sur les relations entre les unités textuelles créées et organisées par les lecteurs.

Les vues sur ces données peuvent être réglées aux besoins d'une pratique de lecture spécifique. A priori un utilisateur a une vue totale sur les anagnoses des autres utilisateurs et sur leurs interprétations et parcours interprétatifs. Il peut ainsi, selon ses besoins, intégrer les résultats interprétatifs d'un autre lecteur et d'une autre anagnose.

Sur la partie droite de la figure, l'interface accompagne l'utilisateur pour la création d'une nouvelle classe sémantique, qui dans ce cas est une isotopie. L'interface guide la justification des attributions du sème et propose des classes pouvant servir à cette justification (en respectant les contraintes contextuelles que nous avons mentionnées dans les sections précédentes).



Figure 4 : Interface utilisateur proposant, à gauche, plusieurs vues sur les mêmes données et à droite, le protocole de création d'une nouvelle interprétation et de sa connexion à ses interprétants.

Les fonctionnalités de projection et de sélection des données offertes par un outil de gestion classique permet également une navigation aisée d'une structure par ailleurs complexe. Il est ainsi possible de demander au programme d'identifier les liens sémantiques ou les similarités entre deux textes analysés, voire de caractériser leurs différences, ou encore de structurer un corpus sur la base des structures

sémantiques repérées dans les textes. Les possibilités ici sont nombreuses et plutôt que de les énumérer au sein d'une interface générique il est en effet préférable de les limiter en les adaptant à des pratiques spécifiques.

4.3. Vérifications des contraintes

Afin de remplir un rôle d'incitation, la machine peut activer un certain nombre de contraintes portant sur les relations déclarées dans la base de données. Nous avons esquissé précédemment certaines de ces contraintes, mais leur définition n'est ni fermée ni absolue : à certains types de parcours sont applicables certaines contraintes.

Rappelons, par exemple :

— La nécessité de « justifier » les attributions de sèmes correspondant à des afférences contextuelles, en exigeant la présence d'une classe sémantique déjà établie dans un contexte proche (même texte ou même anagnose) ou bien d'un sème dans un contexte syntagmatique en tant que lexie.

— La nécessité d'obtenir un réseau de relations (sèmes spécifiques) suffisamment dense pour étayer la cohérence d'une classe sémantique. Ceci se traduit formellement par un ensemble de propriétés à atteindre pour le graphe d'oppositions dont un exemple est présenté dans la Figure 1.

Notons toutefois que ces contraintes sont en quelque sorte « aveugles » quant aux significations et aux sèmes qu'elles régissent. Leur expression est purement formelle et combinatoire. Bien que provenant d'une limite naturelle de l'automatisation de notions sémantiques, elles ont l'avantage de proposer à l'utilisateur un angle nouveau sur des données lourdes de sens. Par exemple, l'exploration d'une classe sémantique sous de telles contraintes peut avoir l'effet bénéfique de rapprocher deux signifiés en demandant à l'utilisateur de caractériser leur différence par un sème, et ainsi d'offrir la possibilité d'un enrichissement. Le rôle de toutes ces contraintes n'est d'ailleurs autre que celui d'enrichir une structure sémantique existante.

4.4. Scénarios d'utilisation

L'outil reste en général malléable. Les contraintes précédemment citées correspondent à certains types de parcours interprétatifs, ou du moins à certaines étapes de ces parcours. Nous explicitons ici deux grands types d'exploration intertextuelle de corpus, dont la dynamique générale va, dans le premier cas, vers l'explicitation de certains phénomènes d'un texte en faisant appel à l'intertexte, et dans le second cas vers la justification de certaines hypothèses intertextuelles.

Pour ce faire, nous devons tout d'abord envisager deux relations principales entre un texte et l'intertexte auquel il est intégré.

4.4.1. Statut du texte dans l'intertexte

Le premier cas est celui du texte premier sur lequel se concentre une approche interprétative. Un certain nombre d'hypothèses peuvent être formulées quant à sa structure sémantique, et confrontées à celle d'autres textes. Le processus principal sera celui de l'établissement de structures locales, dont la justification va servir de moteur à la construction de l'intertexte, et d'une anagnose orientée par ces hypothèses. La même orientation vaut aussi pour les thématiques étudiées.

Dans ce cas, le parcours interprétatif prend effectivement, sinon ontologiquement, racine dans une première exploration des structures textuelles, et la mise en place de schémas de relations sémantiques qui, quoiqu'ancrés sur l'axe syntagmatique, n'en sont pas moins capables d'une abstraction propre à une immersion dans l'intertexte. Un exemple de ce type d'approche est la définition d'un thème dans une œuvre, et l'identification actualisée de ses composantes, comme un schéma actanciel. Cette structure étant établie, elle peut ensuite être confrontée à d'autres textes, et y chercher confirmation, infirmation ou enrichissement.

Le second cas est au contraire celui dont le point de départ est déjà un intertexte constitué. Ou plutôt, un corpus établi autour duquel certaines relations intertextuelles sont préétablies. Ces relations intertextuelles étant soit ancrées, soit des structures actanciennes restant à associer à des unités de surface, elles sont capables d'une projection dans un texte. À partir de cette projection, leur immersion dans la textualité permet à l'interprète, comme on le verra, de les spécifier ou

de les étendre, et de donner par là même naissance à de nouvelles structures connexes ou transversales aux initiales.

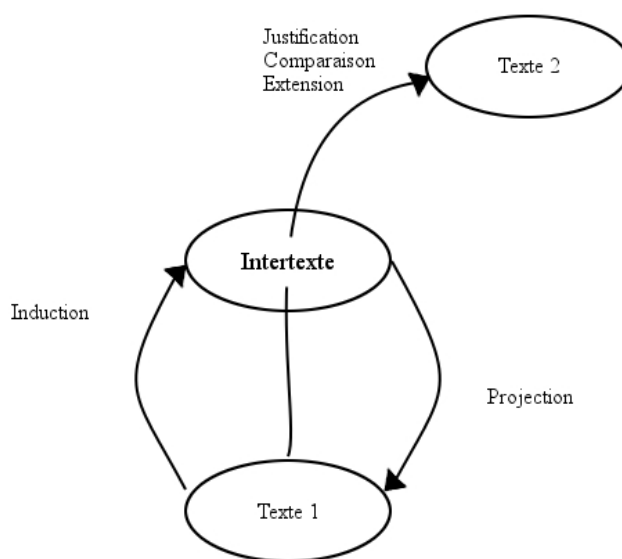


Figure 5 : Relations entre texte et intertexte

Comme il est exprimé dans la Figure 5, les relations orientées entre le texte et l'intertexte peuvent être caractérisées de façon très générale par le principe de projection et d'ancrage contextuel dans un cas, et d'extension, de généralisation ou d'induction dans l'autre sens. Ces deux types de mouvements sont toutefois intrinsèquement liés dans la dynamique d'un parcours interprétatif complet.

Ainsi, notre formalisation des rapports intertextuels se base sur deux types de relations. La première est une extension, en dehors des frontières du texte, des principes d'attribution de traits sémantiques et de la notion d'isotopie. La seconde concerne à un niveau d'abstraction plus élevé, les relations entre textes eux-mêmes. C'est ici que peuvent

s'inscrire les notions classiques des théories de l'intertextualité (parodie, imitation, pastiche), ou même des opérations de structuration de corpus (même auteur, chronologie, genre, etc.).

Ces deux types de relations intertextuelles se répondent et se complètent, une isotopie commune étant une explication ou un indice d'une similarité globale entre deux textes.

4.4.2. Premier cas : du texte à l'intertexte

1. Définition d'une anagnose, et intégration du texte central. Cette déclaration, purement formelle, correspond à l'initialisation de la base de données. L'anagnose initiale ne comprend donc qu'un texte non analysé.

2. Exploration du texte. Identification d'isotopies, formation de classes. Le lecteur utilise l'interface afin d'identifier les lexies et leurs sèmes associés. Cette première partie de l'analyse est sans contraintes, et sert à situer l'analyse au niveau général et intuitif d'une première approche du texte. Ce faisant, un premier ensemble de sèmes est créé, et des lexies sont extraites du texte, et leurs correspondants sémantiques que sont les sémèmes.

Exemple : déclaration de la classe /Animal/ dans les Chants de Maldoror, et repérage des occurrences qui s'y rapportent.

3. Activation de contraintes internes : Identification de sèmes spécifiques. Pour chaque classe formée à l'étape précédente, le lecteur se voit proposer des couples de sémèmes à décrire (par des sèmes) afin de les distinguer. Se forment ainsi des isotopies spécifiques liées au texte.

Exemple : déclaration des connotations /agressivité/ ou /négativité/ au sein de la classe précédente, *i.e.* comme sème spécifique entre les éléments de la classe.

4. Activation des contraintes externes : Identification des attributions à justifier, provenant des sèmes afférents. Le choix peut se faire par des contraintes formelles gérées automatiquement (classes insuffisamment étendues, ou attribution de sèmes isolés), ou par l'utilisateur. Dans le cas d'une contrainte automatique, le lecteur se verra demander par le programme une extension intertextuelle d'une isotopie concernant trop peu d'unités, par exemple.

Exemple : l'isotopie /agressivité/ est ainsi déclarée comme relevant d'une afférence.

5. Induction des isotopies résultant des trois dernières étapes en isotopies intertextuelles. Les isotopies repérées dans le premier texte sont généralisées sur un plan purement formel, et ainsi extensibles à d'autres textes.

6. Extension de l'intertexte : Ajout de texte(s) secondaire(s) au choix du lecteur. Ceci entraîne également la définition dans l'anagnose d'une relation entre textes.

Exemple : L'utilisateur identifie un second texte dans lequel sont présentes des occurrences des sèmes identifiés précédemment, par exemple 'pou'.

7. Projection des isotopies intertextuelles sur le texte secondaire, entraînant la création de classes sémantiques au sein du nouveau texte. A ce stade, ces sèmes ne sont pas encore associés à des signifiés du second texte, seule la relation entre les deux textes est ainsi établie.

8. Ancrage de ces isotopies sur le texte secondaire, repérage de lexies par le lecteur. C'est à cette étape que la source des afférences intertextuelles est établie dans le second texte, afin de satisfaire aux contraintes de l'étape 4.

9. Eventuellement, exploration des nouvelles classes dans le texte secondaire (étape 3), *ad libitum*.

Exemple : L'utilisateur peut par exemple explorer les autres informations liées aux animaux dans le second texte.

Une boucle se met ainsi en place, au sein de chaque texte, par des passages successifs d'une textualité à l'autre ; ceci entraînant la création d'une forme de décor intertextuel autour du texte initial.

4.4.3. Deuxième cas : de l'intertexte au texte

1. Définition d'une anagnose et intégration simultanée d'un ensemble de textes. Comme précédemment, ceci est purement formel et l'anagnose ne contient pas encore de relations sémantiques.

2. Définition de relations entre textes. Par exemple une relation de compatibilité entre un texte et son commentaire, ou entre une *imitation* (cf. Genette 82) et le texte d'origine. Dans ce deuxième cas, ce sont les relations générales entre textes qui sont établies les premières.

3. Définition d'isotopies intertextuelles : Construction d'un schéma sémantique afin de préciser une hypothèse inter-textuelle précédemment définie. Par exemple, identification de thèmes (sèmes) communs à

deux textes, ou de thématiques divergentes ou incompatibles. A ce stade, seuls les sèmes sont déclarés, et ne sont pas encore associés à des sémèmes.

4. Projection et création des classes : Pour chacun des sèmes déclarés précédemment, l'utilisateur procède à l'attribution vers des sémèmes de chaque texte, et crée ainsi des classes sémantiques et des isotopies.

5. Spécification : Au sein de chaque texte et pour chaque classe, explicitation des sèmes spécifiques, de façon à satisfaire les contraintes d'identité sémantique (comme dans le parcours précédent). Cette étape fait donc apparaître de nouvelles isotopies dans chaque texte

6. Eventuellement, ces nouvelles isotopies peuvent, dans certains cas, se retrouver dans les deux textes ce qui renforce les liens sémantiques intertextuels.

Il est à noter que le découpage que nous venons de présenter est effectué pour des raisons de clarté de notre exposé. Les deux chemins font partie d'un parcours interprétatif global qui ressemble plus à une spirale pouvant emprunter à tout moment l'une ou l'autre direction.

5. Conclusion et perspectives

Nous avons donc présenté ici comment un certain type d'interprétation peut être manipulé formellement et informatiquement. Par le biais d'un ensemble de relations entre des unités de surface (texte ou unités lexicales) et des unités de description (sème), et en imposant une organisation de ces relations, il est possible de capter un certain nombre des notions repérées par un lecteur dans un texte ou un ensemble de textes. Plus précisément, ce type d'organisation, tout en restant souple par la nature et le type de ces marqueurs, permet plus avant une collaboration entre un système formel et les intuitions d'un interprète. Nous proposons en effet une forme d'explicitation et de justification de ces intuitions traduites formellement qui va dans le sens d'une exploration sémantique d'un ensemble de textes. Dès lors, la notion d'intertexte prend une place centrale dans notre approche : au-delà d'un simple corpus de travail, l'intertexte devient une construction dynamique, dont les éléments se mettent en place au long d'un parcours.

Devenant ainsi le lieu où se déclarent et s'organisent les relations sémantiques, l'intertexte devient une anagnose, ou une structure sémantique intertextuelle représentant les parcours interprétatifs d'un lecteur selon un objectif donné. Représentée par une base de données de relations entre les différentes unités décrites par des marqueurs sémantiques (textes, lexies, ou relations entre ces unités), elle nécessite une gestion informatique, réalisée par un outil spécifique.

De plus, la construction d'une telle structure sémantique intertextuelle peut et doit être régie par un ensemble de contraintes. Ces contraintes, qui ne reposent pas directement sur le contenu des relations ou des marqueurs sémantiques déclarés par l'utilisateur, à qui nous laissons la liberté et la possibilité d'exprimer sa subjectivité, servent au contraire à le guider vers une explicitation et une forme de rationalisation de ses intuitions. Traduites par un protocole de mise en place de relations sémantiques, ces contraintes deviennent les éléments d'un dialogue entre la machine, qui cherche à construire une base de données cohérentes, et l'utilisateur qui se voit ainsi demandé de déclarer de nouvelles relations, ou d'en expliciter certaines. Les réponses fournies viennent ainsi enrichir autant les descriptions du corpus initial que ce corpus lui-même. L'intertexte devient alors une construction dynamique, dont l'extension est guidée par ces exigences formelles et les découvertes de l'interprète.

Nous avons proposé ici quelques moyens de capter par ce biais certains types d'intertextes ou de relations intertextuelles. Dès lors, la voie est ouverte à y ajouter de nouvelles notions venant des théories de l'intertexte (Genette, 1982). Ainsi, il est envisageable de proposer certaines formes de calcul de relations prédéfinies entre des textes décrits sémantiquement. La déclaration d'une relation de parodie, ou d'imitation, pourrait ainsi être traduite par la nécessité d'atteindre un certain nombre de structures thématiques communes aux deux textes, et bien sûr d'en caractériser en même temps les différences. La nécessité apparaît alors de disposer, pour la description des textes, de structures plus complexes que de simples isotopies, inspirées par exemple des schémas actanciels, comme ceux proposés par F. Rastier (Rastier 95). Bien que basée sur le repérage de sèmes dans un texte, de telles structures nécessitent encore un autre niveau de formalisation.

BIBLIOGRAPHIE

- Assadi H. 1998. *Construction d'ontologies à partir de textes techniques — Application aux systèmes documentaires*. Thèse de Doctorat. Université de Paris 6.
- Genette G. 1982. *Palimpsestes — La littérature au second degré*. Editions du Seuil.
- Habert B., Nazarenko A. et Salem A. 1997. *Les linguistiques de corpus*. Armand Colin.
- Philip M. 1971. *Lectures de Lautréamont*. Armand Colin.
- Pichon R. et Sébillot P. 1999. *Différencier le sens des mots à l'aide du contexte de leur occurrence*. Actes de TALN 99.
- Plett H.F. 1991 (ed.). *Intertextuality*. Walter de Gruyter, Berlin/NY.
- Rastier F. 1987. *Sémantique Interprétative*. PUF.
- Rastier F. 1989. *Sens et Textualité*. Hachette.
- Rastier F., Cavazza M. et Abeille A. 1994. *Sémantique pour l'analyse : de la linguistique à l'informatique*. Masson.
- Rastier F. (éd) 1995. *L'analyse thématique des données textuelles — L'exemple des sentiments*, Paris, Didier.
- Ricœur P. 1986. *Du texte à l'action — Essais d'herméneutique, II*. Editions du Seuil.
- Tanguy L., Armstrong S. et Walker D. 1999. *Isotopies sémantiques et vérification de traduction*. Actes de TALN 99.
- Tanguy L. 1997. *Traitement automatique de la langue naturelle et interprétation : contribution à l'élaboration d'un modèle informatique de la sémantique interprétative*. Thèse de Doctorat — Université de Rennes 1.
- Tanguy L. et Thlivitis Th. 1996. *PASTEL : un protocole informatisé d'aide à l'interprétation des textes*, actes de ILN 96, Nantes.
- Morgan Th. 1985. *Is there an intertext in this text ? Literary and Interdisciplinary Approaches to Intertextuality*. American Journal of Semiotics, 3, 1-40.
- Thlivitis Th. 1998. *Sémantique interprétative intertextuelle : assistance informatique anthropocentrée à la compréhension des textes*. Thèse de Doctorat — Université de Rennes 1.